ELSEVIER

Advances in maize genomics: the emergence of positional cloning

Esteban Bortiri[1], Dave Jackson[2] and Sarah Hake[1]

Positional cloning has been and remains a powerful method for gene identification in *Arabidopsis*. With the completion of the rice genome sequence, positional cloning in rice also took off, including the cloning of several quantitative trait loci. Positional cloning in cereals such as maize whose genomes are much larger than that of rice was considered near impossible because of the vast amounts of repetitive DNA. However, conservation of synteny across the cereal genomes, in combination with new maize resources, has now made positional cloning in maize feasible. In fact, a chromosomal walk is usually much faster than the more traditional method of gene isolation in maize by transposon tagging.

Addresses
[1] Plant Gene Expression Center, USDA-ARS, and Plant and Microbial Biology Department, University of California, Berkeley, 800 Buchanan Avenue, Albany, California 94710, USA
[2] Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA

Corresponding author: Hake, Sarah (maizesh@nature.berkeley.edu)

Current Opinion in Plant Biology 2006, **9**:164–171

This review comes from a themed issue on
Genome studies and molecular genetics
Edited by Susan R Wessler

Available online 2nd February 2006

1369-5266/$ – see front matter

DOI 10.1016/j.pbi.2006.01.006

Introduction

The genetics of maize has a long history, starting with mutants that were identified in the early 1900s [1,2]. Its large chromosomes and chromosomal features, such as knobs, have allowed a rich integration of cytogenetics and genetics [3]. Separate male and female flowers facilitate pollinations and so make the generation of mapping populations easy. Thus, maize has had a robust genetic map for over 70 years. With the advent of transposon tagging and the discovery of additional transposon systems, such as *Mutator* (*Mu*), maize led the way in gene isolation for several years [4]. However, not every gene is easily cloned by transposons, and alleles that arise from transposon screens might not be tagged with the transposon of the parental line. Using unstable alleles [5,6••] makes transposon tagging more likely to succeed.

The first gene to be cloned on the basis of its physical position in crop plants was the resistance to *Pseudomonas syringae pv. tomato* locus (*Pto*) of tomato [7]. The ability to clone this gene came from the time-consuming development of introgressed genetic stocks and extensive marker development. The cloning of the first few genes in *Arabidopsis* also took years, as the workers walked along the chromosome one step at a time. The sequenced *Arabidopsis* genome led to a flourish of gene isolation, such that any *Arabidopsis* gene that has a mutant phenotype is now likely to be cloned multiple times. The knowledge of *Arabidopsis* gene function facilitates our understanding of other plants, although in distantly related species such as maize the gene orthology and conservation of function is often unclear.

Several resources have recently come together to make maize, which is rich with multiple alleles of many mutants, accessible to positional cloning. Here, we review positional cloning strategies in this species as a guide for the success of others, and describe the maize and wheat genes that have recently been cloned by their map position. Maize now stands in a good position for rapid expansion in gene identification.

The maize genome from a positional cloning perspective

Maize has long been known for its high level of polymorphism within coding regions [8]. Recent work suggests that the degree of polymorphism between maize lines is even greater than that between humans and chimpanzees (ES Buckler, pers. comm; see review by Buckler, Gaut and McMullen, this issue). The discovery of vast tracts of retrotransposable elements between genes [9] affirmed maize geneticists' focus on genetics, rather than genomics, and suggested that each inbred might have unique arrangements of retroelements. Sequencing strategies that use gene enrichment have been used successfully to avoid this repetitive DNA [10,11] and now most maize genes have been sequenced to some extent, although not anchored on the genetic map. Two papers have increased our awareness of polymorphism, reporting findings that the 'gene content' differed between maize inbreds [12,13]. It was not clear at that time, however, whether genes were missing in one inbred or inserted in another. A large effort to sequence allelic chromosomal regions in B73 and Mo17 inbreds confirmed these earlier findings and showed that indeed more than 50% of the DNA at any given locus is not co-linear between inbreds [14••]. It appears the non-colinear genes are due to insertion rather than deletion, are members of gene families, and are mostly partial genes or pseudogenes. These gene fragments are thought to be carried by *Helitron* transposons [15•,16•]. This new class

of elements has been identified in *Caenorhabditis elegans* and *Arabidopsis*, although it was missed in earlier transposon prediction programs because of their lack of inverted terminal repeats.

How does this extensive polymorphism between inbreds affect a positional cloning strategy? Given that the *Helitron*-captured genes are predominantly non-functional, and are duplicated in other parts of the genome, their absence would be unlikely to lead to a loss-of-function phenotype. In addition, the genes that are co-linear between maize inbreds are almost always co-linear with rice [14••]. Thus, a positional cloning strategy that relies on rice as a scaffold is not likely to be affected by inbred variation. It remains a possibility that a sought-after gene, especially if defined by a dominant mutation, could be a unique enhancer that is found only in a particular inbred. A sure method to avoid such a scenario is to identify alleles in more than one inbred – if at all possible. In the end, the power of forward genetics should prevail over the complex nature of the maize genome, and the high level of polymorphism will ensure successful marker development to identify most genetically defined loci.

A primer for positional cloning

Defining a rough map position

The principle behind positional cloning resides in the fact that recombination and physical distance are directly related; more recombination events between homologous chromosomes are possible as the distance increases. To observe recombination events in a mapping population, the parents of that population need to differ at as many loci as possible. Fortunately, most maize inbreds are highly polymorphic, as shown by a brute-force analysis of 94 loci across 260 inbreds [17]. Ideally, a mutant will be isolated in a particular inbred line and crossed to a different inbred that carries the wildtype allele. Crossing a mutant to a couple of different inbreds is a good idea to ensure that adequate polymorphisms will be found. B73 and Mo17 are useful inbreds, especially as they are the parents of the largest recombinant inbred population, the intermated B73 x Mo17 population (IBM) [18] for which many markers have been developed.

A first step toward positional cloning is to identify the chromosome arm where your mutation is located. Genetic crosses to translocation stocks can locate a mutation to a chromosome arm in one to two crosses [19,20], but bulk segregant mapping is probably more efficient and quicker [21]. Maize chromosomes are sub-divided into 8–12 subregions, or 'bins' [22]. We have mapped mutations to a bin using DNA from as few as 10 normal plants and 10 mutants from a segregating population, and two markers per chromosome arm. Once a mutant is placed on a chromosome arm using bulk segregants, DNA preparations [23] from individuals are used to map the mutation to a defined interval. The number of plants that is required varies, but 1000 is probably the lower limit to ensure the mapping of any mutation to a single bacterial artificial chromosome (BAC). This number is in the same magnitude as that required in *Arabidopsis* [24], even though the maize genome is about 20 times larger than that of *Arabidopsis*. In retrospect, this finding is not too surprising, as recombination occurs mostly within genes [25], and the gene numbers in *Arabidopsis* and maize might not be too different [26].

To start the walk, distal and proximal flanking markers that are PCR-based and co-dominant should first be identified. They do not need to be the closest markers but must be reliable. Recombinants that show exchange of flanking markers are identified, and these individuals are grown to maturity to score the phenotype and to make larger high-quality DNA preps. These recombinants are extremely important and will be used to refine the position.

Useful maps and integrated markers can be found at a couple of different sites. IBM maps, assembled using data from the Maize Mapping Project (http://www.maizemap.org/), are periodically updated with the latest versions available at the MaizeGDB website (www.maizegdb.org, go to 'Maps'). The latest version of the IBM map combines 3149 gene-based insertion/deletion polymorphisms (IDPs), which are PCR-based polymorphisms between B73 and Mo17, and 2030 markers from the Missouri Mapping Project. The list of IDPs, the sequence from which they are derived, and their map position can be obtained from the Maize Genetic Mapping Project website (http://maize-mapping.plantgenomics.iastate.edu/). These web pages provide primer sequence information and, in some cases, PCR conditions.

Sources of additional markers

As one increases the number of individuals in the mapping population, it is important to identify additional markers to narrow the mapping interval. Until recently, this step was the limiting factor and the main obstacle to positional cloning in maize. Now, several sequencing projects have provided a wealth of information about gene islands in the maize genome [27•]. In the following paragraphs, we outline the two main sources of information to identify low-copy number genes that can be used as markers in a selected genomic region, and we provide the URLs of some of the *in-silico* resources from which the information can be retrieved (Table 1). The steps described here are meant to be suggestions and are by no means the only way to retrieve the information.

BAC-anchored genetic markers and ESTs

In addition to providing a physical map of the maize genome, the BAC contigs provide a source of markers. They have been probed with gene-specific oligonucleotide-based ('overgo') probes designed from expressed

Table 1

Useful resources for positional cloning in maize.

Name	URL	Main tools
MaizeGDB	www.maizegdb.org	Genetic maps, markers information and EST contigs database.
Arizona Genomics Institute	www.genome.arizona.edu/fpc/maize	BAC contigs and BAC-end sequences.
Gramene	www.gramene.org	Rice genome annotation, grass comparative genomics.
MAGI	www.plantgenomics.iastate.edu/maize	GSS contigs database.
TIGR Rice Genome Annotation	rice.tigr.org/	Starting place for rice analysis at TIGR. Users can carry out blast or use a genome browser.
TIGR Maize Database	Maize.tigr.org/	Starting place for analysis of maize.
TIGR Maize Gene Index	www.tigr.org/tigr-scripts/tgi/T_index.cgi?species=maize	EST and GSS (AZM) contigs database.
TIGR Maize Genomic Blast Search	tigrblast.tigr.org/tgi_maize/index.cgi	Blast sequences against maize and sorghum assemblies.
TIGR Maize Marker Mapping	maize.tigr.org/tigr-scripts/tgi/marker2.annotator.pl?species=combo_marker	Find markers in maize sequence assemblies placed on the genetic map.
Maize Genetic Mapping Project	maize-mapping.plantgenomics.iastate.edu	IDP markers.
Maize Mapping Project	www.maizemap.org	Links genetic and physical maps.
Maize Seq	www.maizeseq.org	EST contigs database (proprietary).

sequence tag (EST) sequences [28]. This BAC-overgo detection allows the anchoring of genes to the physical map, thereby providing a link between the physical and the genetic map. Another advantage of this BAC marker alignment is that it provides a potential source of additional closely linked markers. The Arizona Genomics Institute Web FPC maps (http://www.genome.arizona.edu/fpc_hicf/maize/) can be searched by marker (for example, from the IBM map) or by BAC clone number. Once a contig is displayed, a toggle in the upper right corner allows one to highlight overgo markers that are derived from maize ESTs that have a clear ortholog in rice (select 'Gramene'), or to obtain sequence information for a particular overgo (select 'GenBank'). A third toggle links some overgos to the MaizeGDB website. The advantage of using sequence information from BAC contigs to develop new polymorphic markers is that they are very likely to be linked to the gene in question. A caveat is that some BAC-overgo detections are promiscuous. One way to minimize the possibility of landing in the wrong genomic area is to avoid overgos that detect BACs that are placed in many contigs. These potential artifacts probably arise from overgos that detect members of gene families or from inaccurate BAC placement. Another source of markers that are anchored to BACs are BAC-end sequences (BESs) [29]. A total of 154 293 BESs generated by the Messing group from the ZMMBBb B73 library are available at the same website as the BAC contigs.

Rice genes as queries in maize databases and vice versa

Despite having diverged several million years ago, the genomes of cereals show considerable conservation in gene number and order [30]. Because of this conserved synteny, the rice genome can be used as a guide to find genetically linked markers in maize. A first step is to use sequences from which IBM markers have been developed to 'blast' the rice genome databases. The Gramene website is the most useful resource for this purpose (http://www.gramene.org/). Ideally, all maize markers that are linked to the gene under investigation should locate to the same region of the rice genome. It is very common, however, to find small translocations between rice and maize. Once a rice region that is syntenous to the maize flanking interval has been identified, the additional rice genes in this region can be used to find their orthologous maize sequences by either blasting them in maize databases or using the comparative map (cMap) feature of Gramene. These pages display the annotated rice pseudomolecules and the grass sequences that match specific regions of the rice chromosome. Not all maize ESTs and genome survey sequences (GSS) will be present in this comparative map, however, and thus manual blast searches are a useful way to find additional markers.

Linking the physical and genetic maps

The assembly of maize BAC clones into contigs has improved dramatically in the past year thanks to fluorescence-based high-information content fingerprinting (HICF) techniques [31•]. It is likely that a distance of 1 cM is contained within an average-sized BAC contig and that a 0.1 cM distance is contained on 1–2 BACs. If no BAC or BAC contig is available for a particular marker, or if one wishes to assemble a contig, a BAC library needs to be screened. The most widely used BAC library, ZMMBBb, consists of 247 680 clones, with an average insert size of 137 kb, that cover 14 genome equivalents of the B73 inbred line. It is available from the Clemson University Genomics Institute (http://www.genome.clemson.edu/). Another BAC-library resource, ZMMBBc, is available from Children's Hospital in Oakland Research Institute (CHORI; http://bacpac.chori.org/maize201.htm). Once the mapping interval is small enough that it spans a few BACs and a few genes in the orthologous rice region, the first, and easiest, step is to look for a

likely candidate among the rice genes and to identify its maize equivalent. Sequencing the maize BACs provides another way to generate additional markers and/or candidate genes. The low cost of sequencing and the use of computer routines to blast sequences automatically makes sequencing a cost-effective approach to gene discovery in the defined interval. Once candidate genes are identified by recombination and sequencing, alleles of the mutant locus should be sequenced to prove that the correct gene has been identified. If only one allele is available, additional alleles should be generated by forward or reverse genetic approaches [32–34] to prove that the correct gene has been isolated, and even transformation can be used for confirmation [35].

Gene isolation

In the following section, we describe some recent examples of genes that have been cloned by positional cloning or, if they provide important information about maize genome organization, by traditional transposon tagging approaches.

Isolation of quantitative trait loci

The first published papers to use positional cloning in maize and wheat identified quantitative trait loci (QTL). No doubt, this strategy was taken because QTL are defined by their position and are not amenable to transposon tagging. In each case, large numbers of individuals were used to define a small interval that contained the gene.

Doebley and Stec [36] mapped the differences between maize and teosinte to five major QTL. One of these QTL mapped to the region containing *teosinte branched1* (*tb1*), whose mutant phenotype suggests that it is an excellent candidate gene. Cloning of the gene and analysis of both maize and teosinte alleles has supported *tb1* as the gene for this QTL [37,38]. *tb1* was cloned using the traditional method of transposon tagging because a candidate maize mutation was available. However, another QTL that affected the hardness of the seed coat, *teosinte glume architecture* (*tga1*), did not map to a candidate gene. Crosses of the maize *tga1* allele into teosinte and the teosinte *tga1* allele into maize showed that this gene behaves as a single Mendelian locus [39]. Using flanking markers, Doebley and colleagues [39] introgressed the region from teosinte into maize for six generations. They then set up mapping populations and obtained a tightly linked marker, which was used to identify a BAC contig. The BAC contig, along with rice synteny, helped these authors to develop additional markers. Given that Doebley and colleagues couldn't rely on multiple alleles to prove that the gene was cloned, they used the power of recombination to narrow the QTL region to just 1042 base pairs using 3106 individuals. This region was sequenced in 16 maize and 12 teosinte lines. Seven fixed differences were found, six in the promoter, and one in the coding sequence that resulted in the substitution of a lysine in teosinte for an arginine in maize. This single difference in coding sequence is thought to be causative, as there are no obvious RNA expression differences but the teosinte protein is more abundant. Furthermore, an ethyl methyl sulfonate (EMS) revertant of *tga1* had an amino-acid substitution in the position adjacent to the maize arginine [40••]. It is impressive that the first publication of positional cloning in maize should be a QTL, a true *tour d'force*.

Given the complexity of the wheat genome, the cloning of two QTLs in wheat is even more remarkable. Vernalization in wheat is controlled by the *VERNALIZATION1* (*VRN1*) and *VRN2* genes [41,42]. Positional cloning of the *VRN1* gene was carried out using 6190 gametes and a detailed genetic map of wheat [43]. Almost perfect micro co-linearity in this region between wheat, rice and sorghum helped to provide additional markers and to narrow down candidate genes. No recombinants were found for two closely linked rice genes, *APETALA1* and *AGLG1*, which are both MADS-box genes. Expression data strongly supported the hypothesis that the wheat *AP1* is *VRN1* as this gene is upregulated in leaves in response to weeks of cold treatment. *AGLG1* is also expressed in spikes, but is not affected by cold treatment. Sequencing of *AGLG1* in wheat and spring accessions showed no differences, whereas a deletion was found in the promoter of the spring variety of *AP1*, again supporting the idea that *AP1* is *VRN1*.

VRN2 was cloned using a population of 5698 gametes, which defined a 0.04 cM interval [44••]. Yan *et al.* [44••] sequenced this interval (438 kb) and found eight genes. The region was also co-linear with rice and barley, revealing that a 7 kb distance in rice translated to 328 kb in wheat. Additional markers allowed Yan *et al.* to further delimit a region containing three genes. The best *Vrn2* candidate gene, *ZCCT1*, shows an opposite expression pattern to *VRN1*: its level decreased in apices during vernalization. Further support that *ZCCT1* is *VRN2* came from sequencing in winter and spring wheat varieties, in which only this gene showed any polymorphism, an arginine in winter and a tryptophan in spring wheat. All winter wheats have an arginine in the equivalent position, whereas spring wheats have the tryptophan or carry deletions of the gene. Supporting evidence was also provided by the use of RNA interference to modulate *VRN2* levels and thereby change flowering time [44••].

In summary, the use of large mapping populations and synteny with other cereals provided one to a small number of candidate QTL genes. Sequencing and expression analysis across naturally occurring alleles allowed them to identify the most likely candidate. With the newly emerging maize resources, other well-mapped QTL, such as *VGT1* [45], will be cloned soon.

Isolation of genes that regulate inflorescence architecture

Identifying a developmental gene by positional cloning has several advantages over the isolation of a QTL: phenotyping is usually more reproducible; multiple alleles are likely to exist; one can choose the inbred to develop the mapping population; and many developmental genes are conserved across species, making a candidate-gene approach very appealing. In addition, mapping populations can be relatively small (Table 2).

Several genes that regulate inflorescence architecture in maize have been cloned in the past year. Maize tassels have long branches at their base, whereas ears do not. However, both ears and tassels are covered by short branches, called spikelet pairs. Spikelets are found in all grasses, but the arrangement of spikelets into pairs is only found in the *Andropogoneae*, a subgroup of the grasses that includes sorghum, sugar cane, maize and about 1000 other wild grasses [46]. *ramosa1* (*ra1*) mutants have a highly branched tassel and a branched ear. The spikelet pairs of this mutant are replaced by a gradient of indeterminate branches, giving it a Christmas-tree-like appearance. Fortunately, Vollbrecht *et al.* [6••] set out to clone *ra1* by transposon tagging and not using synteny with rice, because the gene, which encodes a zinc-finger transcription factor, is found in maize and sorghum but not rice. *ra1* is expressed at the base of the spikelet-pair meristem. The mutant's phenotype, expression pattern, and absence from a species that does not make spikelet pairs suggest that *ra1* is required for spikelet-pair identity.

Two other maize inflorescence genes have been cloned in the past year using the map position of the mutation in combination with an obvious candidate gene. *barren stalk1* (*ba1*) mutants lack tassel branches and spikelets and are missing ears [47]. The mutation was mapped to a syntenous region in rice that contains a similar mutant phenotype called *lax panicle*. The positional cloning of *lax panicle* [48] provided a candidate gene for *ba1*. One *ba1* allele of this basic helix–loop–helix protein has a *Helitron* upstream of the coding region and a second allele was found by reverse genetics, confirming that the candidate gene is *ba1* [49•]. In a second example, a maize *clavata1* (*clv1*) ortholog was mapped to chromosome 5 in the same region as *thick tassel dwarf1* (*td1*). The phenotype of *td1* mimics that of *Arabidopsis clv* mutants, which have larger inflorescence meristems and more floral organs. Proof that *td1* was the *clv1* ortholog came from analysis of a large number of *Mu*-induced alleles [50•]. The orthologous rice mutation, *FLORAL ORGAN NUMBER1* (*FON1*), was identified at the same time by positional cloning [51].

ra2 [52••], *ra3* and *tasselseed4* (*ts4*) (Figure 1) have also been cloned positionally (D Jackson, G Chuck, S Hake, unpublished). All three of these genes affect the spikelet-pair meristem, but unlike *ra1*, they are found in rice. For the cloning of *ra2*, a population of 1070 individuals was used [52••]. This population narrowed the genetic distance to 0.05 cM on the distal side and 0.8 cM on the proximal side. A BAC contig was identified for the distal side by hybridization with the *mmp186* marker, and this contig turned out to contain *ra2*. The BAC contigs for *mmp186* and *asg48*, the closest proximal marker, did not join and the rice orthologs of *asg48* and *mmp186* were unlinked. Interestingly, *asg48* was the only non-syntenic marker in this region. Two additional markers were obtained within the *mmp186* BAC contig that positioned *ra2* within a smaller interval. A *LATERAL ORGAN BOUNDARY* (*LOB*) rice gene in the syntenic region was completely linked to *ra2* (0 recombinants in 1070 individuals), and indeed, analysis of five mutant alleles (both deletion and insertion alleles) proved that the *LOB* gene was *ra2*.

ra3 was mapped using bulk segregant analysis, and the population was expanded to about 1700 mutant plants.

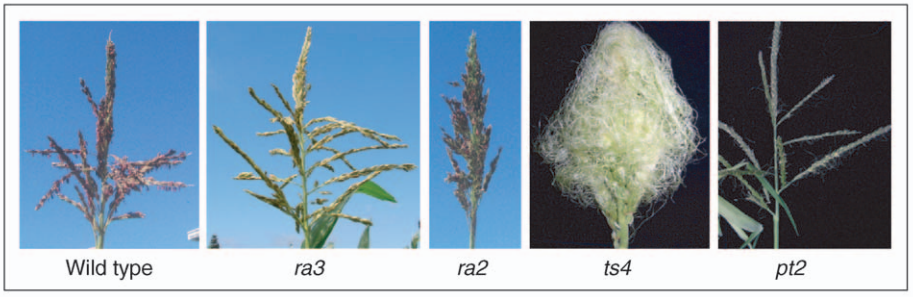
Table 2

Intervals defined by positional cloning.

Gene	Number of individuals[a]	Colinearity with rice	Distance between flanking markers[b]	Distance in rice[c]	Reference
tga1	3105	Yes	1 kb		[40••]
VRN1	6190[d]	Yes	0.03 cM, 324 kb	<10 kb	[43]
VRN2	5698[d]	Yes	0.04 cM, 315 kb	<5 kb	[44••]
ra2	1070	Yes[e]	0.85 cM	65 kb, 11 genes	[52••]
pt2	1030	Yes	1.16 cM	130 kb, 11 genes	L Bartling, H Sakai, S Hake (unpublished)
ra3	1700	Yes	0.2 cM, 6 kb		N Satoh, H Sakai, D Jackson (unpublished)
ts4	1600	Yes	0.3 cM, 1 BAC	110 kb, 10 genes	G Chuck, H Sakai, S Hake (unpublished)
Wab	1400	Yes[e]	10 BACs		H Candela, S Hake (unpublished)
mwp	600	Yes[e]	0.66 cM, 9 BACs		H Candela, A Gerhold, S Hake (unpublished)

[a] Number of individuals.
[b] Refers to the distance in maize or wheat, measured in either cM or number of BACs (average size 130 kb).
[c] Distance in rice between the rice orthologs for the maize flanking markers.
[d] Number of gametes.
[e] Overall synteny was found, but some local rearrangements exist.

Figure 1



The wildtype tassel is compared to four mutants that have increased branching and feminization.

The region was narrowed to 1.6 cM, which corresponded to 350 kb, with the use of additional markers. Sequencing of two BACS revealed seven predicted genes in one BAC and 22 in the other. Additional markers derived from these BACs were used to narrow the *ra3* region to 6 kb or 0.2 cM, which contained a single predicted gene. *ts4* was mapped to a region containing 10 rice genes. Sequencing of maize BAC clones that spanned this region actually found three fewer genes in this interval than in the syntenous rice interval. All seven of these genes were tightly linked to *ts4* (0/1600), thus analysis of RNA levels in multiple alleles was needed to find the correct gene.

These three examples of positional cloning show that rice synteny is a powerful tool for the identification of genes, especially for genes that function in development. The same conservation of synteny might not be as common for other traits, such as disease resistance. For example, the barley stem rust resistance gene, *Rpg1*, was not found in the syntenous region of rice [53]. The growing collection of cloned maize inflorescence genes provides tools to dissect the pathways that lead to tassel and ear development, and to understand morphological diversity in cereal inflorescences.

Conclusions

Positional cloning is a sure-proof method of gene identification. By providing a marker-bound mapping interval, recombination data assure us that we are at a specific genomic region that unequivocally contains the gene of interest. Sooner or later, the gene will be obtained. Positional cloning remains the only way to obtain a gene that underlies a quantitative trait, unless a corresponding mutation can be identified. It is quickly becoming the preferred method of isolating genes that are defined by mutations in maize. The large genome size of maize is not a hindrance, and conserved micro-synteny with rice and improved maize BAC contigs facilitate the effort. Once the maize genome is fully sequenced and assembled, the high frequency of recombination will make it as easy to clone a gene in maize as it is in *Arabidopsis*. Although transformation remains slow, the ease of obtaining additional alleles through targeted tagging or reverse genetics provides proof of gene identification. The diversity that is being catalogued in maize inbreds [54], and the ability to sequence the related teosinte species can also provide a second level of confirmation that the correct gene has been isolated, which is especially useful in the case of QTL-identification projects.

Acknowledgements

We thank George Chuck, Hector Candela-Anton, Linnea Bartling, and Namiko Satoh for sharing their unpublished mapping data. We thank Hector Candela-Anton and Hajime Sakai for valuable comments on the manuscript and the National Science Foundation (NSF; DBI-0110189) for support.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Emerson RA, Beadle GW, Fraser AC: **A summary of linkage studies in maize**. *Cornell Univ Agric Exo Sta Mem* 1935, **180**:1-83.

2. Coe EH Jr, Neuffer MG, Hoisington DA: **The genetics of corn**. In *Corn and Corn Improvement, 3rd Edition*. Edited by Sprague GF, Dudley JW. Am Soc Agronomy Inc, Crop Science Soc Am Inc, Soil Science Soc Am Inc; 1988: 81-258.

3. McClintock B: **Chromosome organization and genic expression**. *Cold Spring Harb Symp Quant Biol* 1951, **16**:13-47.

4. Walbot V: **The *Mutator* transposable element family of maize**. *Genet Eng* 1991, **13**:1-37.

5. Stern DB, Hanson MR, Barkan A: **Genetics and genomics of chloroplast biogenesis: maize as a model system**. *Trends Plant Sci* 2004, **9**:293-301.

6. •• Vollbrecht E, Springer PS, Goh L, Buckler ES, Martienssen R: **Architecture of floral branch systems in maize and related grasses**. *Nature* 2005, **436**:1119-1126.
The classical *ra1* inflorescence-branching mutant of maize was isolated by transposon tagging. *ra1* encodes a transcription factor that is expressed in a boundary domain near the base of meristems and that acts non-autonomously to impose short branch identity as branch meristems are initiated. Of particular interest to genome organization, *ra1* appears to be absent from rice, and variation in its expression might account for the variation in inflorescence architectures across the grasses and throughout maize domestication.

7. Martin GB, Brommonschenkel SH, Chunwongse J, Frary A, Ganal MW, Spivey R, Wu T, Earle ED, Tanksley SD: **Map-based cloning of a protein kinase gene conferring disease resistance in tomato**. *Science* 1993, **262**:1432-1436.

8. Shattuck-Eidens DM, Bell RN, Neuhausen SL, Helentjaris T: **DNA sequence variation within maize and melon: observations from polymerase chain reaction amplification and direct sequencing**. *Genetics* 1990, **126**:207-217.

9. SanMiguel P, Tikhonov A, Jin Y-K, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z *et al.*: **Nested retrotransposons in the intergenic regions of the maize genome**. *Science* 1996, **274**:765-768.

10. Palmer LE, Rabinowicz PD, O'Shaughnessy AL, Balija VS, Nascimento LU, Dike S, de la Bastide M, Martienssen RA, McCombie WR: **Maize genome sequencing by methylation filtration**. *Science* 2003, **302**:2115-2117.

11. Whitelaw CA, Barbazuk WB, Pertea G, Chan AP, Cheung F, Lee Y, Zheng L, van Heeringen S, Karamycheva S, Bennetzen JL *et al.*: **Enrichment of gene-coding sequences in maize by genome filtration**. *Science* 2003, **302**:2118-2120.

12. Fu H, Dooner HK: **Intraspecific violation of genetic colinearity and its implications in maize**. *Proc Natl Acad Sci USA* 2002, **99**:9573-9578.

13. Song R, Messing J: **Gene expression of a gene family in maize based on noncollinear haplotypes**. *Proc Natl Acad Sci USA* 2003, **100**:9055-9060.

14. •• Brunner S, Fengler K, Morgante M, Tingey S, Rafalski A: **Evolution of DNA sequence nonhomologies among maize inbreds**. *Plant Cell* 2005, **17**:343-360.

The authors sequenced more than 2.8 Mb of genomic DNA from several corresponding regions of the B73 and Mo17 genomes. More than 50% of the compared regions were not co-linear. Remarkably, more than one-third of the genes (27/72) were absent in one of the inbreds at these loci; however, the non-shared genes were usually truncated and therefore likely not to be functional.

15. • Lai J, Messing J, Dooner HK: **Gene movement by *Helitron* transposons contributes to the haplotype variability of maize**. *Proc Natl Acad Sci USA* 2005, **102**:9068-9073.

Following on their earlier report of 'intraspecific violation of genetic colinearity' at the *bronze* genomic region of maize [12], the authors show that many of the non-co-linear genes in maize are carried in *Helitron* transposons. Maize *Helitrons* capture genes or gene fragments from several loci and move them around the genome, and so could be important in gene evolution.

16. • Morgante M, Brunner S, Pea G, Fengler K, Zuccolo A, Rafalski A: **Gene duplication and exon shuffling by helitron-like transposons generate intraspecies diversity in maize**. *Nat Genet* 2005, **37**:997-1002.

A comprehensive study of gene content in two maize inbreds revealed as many as 10 000 polymorphisms, mainly generated by DNA insertions. A closer analysis showed that the majority of these were due to *Helitron* rolling-circle transposition. The authors hypothesize that these events could lead to exon shuffling and to the evolution of new proteins.

17. Liu K, Goodman M, Muse S, Smith JS, Buckler E, Doebley J: **Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites**. *Genetics* 2003, **165**:2117-2128.

18. Lee M, Sharopova N, Beavis WD, Grant D, Katt M, Blair D, Hallauer A: **Expanding the genetic map of maize with the intermated B73 x Mo17 (IBM) population**. *Plant Mol Biol* 2002, **48**:453-461.

19. Beckett JB: **Locating recessive genes to chromosome arm with B-A translocations**. In *The Maize Handbook.* Edited by Freeling M, Walbot V: Springer-Verlag; 1993:315-327.

20. Laughnan JR, Gabay-Laughnan S: **The placement of genes using waxy-marked reciprocal translocations**. In *The Maize Handbook.* Edited by Freeling M, Walbot V: Springer-Verlag; 1993: 254-257.

21. Michelmore RW, Paran I, Kesseli RV: **Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations**. *Proc Natl Acad Sci USA* 1991, **88**:9828-9832.

22. Gardiner J, Coe EH, Melia-Hancock S, Hoisington DA, Chao S: **Development of a core RFLP map in maize using an immortalized F2 population**. *Genetics* 1993, **134**:917-930.

23. Konieczny A, Ausubel F: **A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers**. *Plant J* 1993, **4**:403-410.

24. Lukowitz W, Gillmor CS, Scheible WR: **Positional cloning in *Arabidopsis*. Why it feels good to have a genome initiative working for you**. *Plant Physiol* 2000, **123**:795-805.

25. Fu H, Zheng Z, Dooner HK: **Recombination rates between adjacent genic and retrotransposon regions in maize vary by 2 orders of magnitude**. *Proc Natl Acad Sci USA* 2002, **99**:1082-1087.

26. Bennetzen JL, Coleman C, Liu R, Ma J, Ramakrishna W: **Consistent over-estimation of gene number in complex plant genomes**. *Curr Opin Plant Biol* 2004, **7**:732-736.

27. • Fu Y, Emrich SJ, Guo L, Wen TJ, Ashlock DA, Aluru S, Schnable PS: **Quality assessment of maize assembled genomic islands (MAGIs) and large-scale experimental verification of predicted genes**. *Proc Natl Acad Sci USA* 2005, **102**:12282-12287.

The authors report an improved assembly of maize gene-enriched sequences, which they call maize assembled genomic islands (MAGIs). Over 110 000 contigs were assembled, and comparison of a subset to BAC sequences suggests that at least 97% were correctly assembled. The data suggest that the maize genome contains between 33 000 and 54 000 expressed genes.

28. Gardiner J, Schroeder S, Polacco ML, Sanchez-Villeda H, Fang Z, Morgante M, Landewe T, Fengler K, Useche F, Hanafey M *et al.*: **Anchoring 93 971 maize expressed sequence tagged unigenes to the bacterial artificial chromosome contig map by two-dimensional overgo hybridization**. *Plant Physiol* 2004, **134**:1317-1326.

29. Messing J, Bharti AK, Karlowski WM, Gundlach H, Kim HR, Yu Y, Wei F, Fuks G, Soderlund C, Mayer KFX *et al.*: **Sequence composition and genome organization of maize**. *Proc Natl Acad Sci USA* 2004, **101**:14349-14354.

30. Devos KM: **Updating the 'Crop Circle'**. *Curr Opin Plant Biol* 2005, **8**:155-162.

31. • Nelson WM, Bharti AK, Butler E, Wei F, Fuks G, Kim H, Wing RA, Messing J, Soderlund C: **Whole-genome validation of high-information-content fingerprinting**. *Plant Physiol* 2005, **139**:27-38.

Fluorescence-based high-information-content fingerprinting (HICF) techniques are used to assemble BAC clones into contigs. HICF has substantially greater sensitivity and accuracy in forming contigs than fingerprinting using traditional gel-staining methods. This improved physical map should provide an excellent starting point for maize whole-genome sequencing.

32. May BP, Liu H, Vollbrecht E, Senior L, Rabinowicz PD, Roh D, Pan X, Stein L, Freeling M, Alexander D *et al.*: **Maize-targeted mutagenesis: a knockout resource for maize**. *Proc Natl Acad Sci USA* 2003, **100**:11541-11546.

33. Taguchi-Shiobara F, Yuan Z, Hake S, Jackson D: **The *fasciated ear2* gene encodes a leucine rich repeat receptor like protein that regulates shoot meristem proliferation in maize**. *Genes Dev* 2001, **15**:2755-2766.

34. Till BJ, Reynolds SH, Weil C, Springer N, Burtner C, Young K, Bowers E, Codomo CA, Enns LC, Odden AR *et al.*: **Discovery of induced point mutations in maize genes by TILLING**. *BMC Plant Biol* 2004, **4**:12.

35. Magnard JL, Heckel T, Massonneau A, Wisniewski JP, Cordelier S, Lassagne H, Perez P, Dumas C, Rogowsky PM: **Morphogenesis of maize embryos requires *ZmPRPL35-1* encoding a plastid ribosomal protein**. *Plant Physiol* 2004, **134**:649-663.

36. Doebley J, Stec A: **Inheritance of the morphological differences between maize and teosinte: comparison of results for two F2 populations**. *Genetics* 1993, **134**:559-570.

37. Doebley J, Stec A, Hubbard L: **The evolution of apical dominance in maize**. *Nature* 1997, **386**:485-488.

38. Wang R-L, Stec A, Hey J, Lukens L, Doebley J: **The limits of selection during maize domestication**. *Nature* 1999, **398**:236-239.

39. Dorweiler J, Stec A, Kermicle J, Doebley J: ***Teosinte glume architecture1*: a genetic locus controlling a key step in maize evolution**. *Science* 1993, **262**:233-235.

40. •• Wang H, Nussbaum-Wagler T, Li B, Zhao Q, Vigouroux Y, Faller M, Bomblies K, Lukens L, Doebley JF: **The origin of the naked grains of maize**. *Nature* 2005, **436**:714-719.
The authors positionally cloned the *teosinte glume architecture* (*tga1*) locus, which encodes a transcriptional regulator. This locus was selected during maize domestication, resulting in the exposure of the seed from the hard casing that envelops it in its ancestor, teosinte. This work provides another classic example of how variation in a single gene can lead to dramatic phenotypic changes during evolution.

41. Dubcovsky J, Lijavetzky D, Appendino L, Tranquili G: **Comparative RFLP mapping of *Triticum monococcum* genes controlling vernalization requirement**. *Theor Appl Genet* 1998, **97**:968-975.

42. Tranquilli G, Dubcovsky J: **Epistatic interaction between vernalization genes *Vrn-Am1* and *Vrn-Am2* in diploid wheat**. *J Hered* 2000, **91**:304-306.

43. Yan L, Loukoianov A, Tranquili G, Helguera M, Fahima T, Dubcovsky J: **Positional cloning of the wheat vernalization gene *VRN1***. *Proc Natl Acad Sci USA* 2003, **100**:6263-6268.

44. •• Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, SanMiguel P, Bennetzen JL, Echenique V, Dubcovsky J: **The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization**. *Science* 2004, **303**:1640-1644.
VRN2 was the second wheat vernalization gene to be cloned by position. *VRN2* has an expression pattern opposite to that of *VRN1*. Although the wheat vernalization genes have homologs in *Arabidopsis*, they carry out different functions. The barley *VRN2* ortholog also regulates vernalization.

45. Salvi S, Tuberosa R, Chiapparino E, Maccaferri M, Veillet S, van Beuningen L, Isaac P, Edwards K, Phillips RL: **Toward positional cloning of *Vgt1*, a QTL controlling the transition from the vegetative to the reproductive phase in maize**. *Plant Mol Biol* 2002, **48**:601-613.

46. Goup GPW: **A phylogeny of the grass family (Poaceae), as inferred from eight character sets**. In *Proceedings of the Second International Conference on the Comparative Biology of the Monocotyledons. Vol 2. Symposium on Grass Systematics and Evolution.* Edited by Everett JE, Jacobs SWL: CSIRO; 2000.

47. Ritter MK, Padilla CM, Schmidt RJ: **The maize mutant, *barren stalk 1*, is defective in axillary meristem development**. *Am J Bot* 2002, **89**:203-210.

48. Komatsu K, Maekawa M, Ujiie S, Satake Y, Furutani I, Okamoto H, Shimamoto K, Kyozuka J: **LAX and SPA: major regulators of shoot branching in rice**. *Proc Natl Acad Sci USA* 2003, **100**:11765-11770.

49. • Gallavotti A, Zhao Q, Kyozuka J, Meeley RB, Ritter MK, Doebley JF, Pe ME, Schmidt RJ: **The role of *barren stalk1* in the architecture of maize**. *Nature* 2004, **432**:630-635.
A candidate gene approach, together with knowledge of map location, led to the isolation of the maize *barren stalk1* (*ba1*) gene, which is required for the initiation of axillary meristems in maize. Interestingly, *ba1* appears to have been selected during maize domestication, presumably for its effects on shoot and inflorescence architecture.

50. • Bommert PB, Lunde C, Nardmann J, Vollbrecht E, Running MP, Jackson D, Hake S, Werr W: ***thick tassel dwarf1* encodes a putative maize orthologue of the *Arabidopsis* CLAVATA1 leucine-rich receptor-like kinase**. *Development* 2005, **132**:1235-1245.
Maize homologs of the *Arabidopsis CLAVATA1* gene were isolated, and one mapped near the *thick tassel dwarf1* (*td1*) mutation. Characterization of multiple *td1* alleles confirmed that it is a *CLV1*-like gene. Interestingly, *td1* maps in the region of QTLs for seed row number, spikelet density and plant height, suggesting that it might have been a target of selection during maize breeding or domestication.

51. Suzaki T, Sato M, Ashikari M, Miyoshi M, Nagato Y, Hirano HY: **The gene *FLORAL ORGAN NUMBER1* regulates floral meristem size in rice and encodes a leucine-rich repeat receptor kinase orthologous to *Arabidopsis* CLAVATA1**. *Development* 2004, **131**:5649-5657.

52. •• Bortiri E, Chuck G, Vollbrecht E, Rocheford TF, Martienssen R, Hake S: ***ramosa2* encodes a LOB domain protein that determines the fate of stem cells in branch meristems of maize**. *Plant Cell* 2006. in press.
The authors walk to *ra2* and find it encodes a LOB domain-containing protein. *ra2* functions to limit branching in the maize inflorescence and is expressed in the cells that will become the branch. It is conserved in sequence and expression in barley, rice and sorghum.

53. Brueggeman R, Rostoks N, Kudrna D, Kilian A, Han F, Chen J, Druka A, Steffenson B, Kleinhofs A: **The barley stem rust-resistance gene *Rpg1* is a novel disease-resistance gene with homology to receptor kinases**. *Proc Natl Acad Sci USA* 2002, **99**:9328-9333.

54. Buckler E, Thornsberry J: **Plant molecular diversity and applications to genomics**. *Curr Opin Plant Biol* 2002, **5**:107-111.